Persistence and Monotony Properties of Argumentation Semantics

Tjitze Rienstra[1] Chiaki Sakama[2] Leendert van der Torre[1]

[1]University of Luxembourg

[2]Wakayama University

July 24, 2015

Overview

Introduction

- Argumentation is a dynamic process,
  - but an argumentation framework represents only a static snapshot.
- To consider dynamics, we need to consider AFs that change,
  - e.g. addition of new arguments and attacks.
- How do semantics for argumentation behave when the AF changes?
- 2009: Overview for grounded semantics (Boella, Kaci, vdT)
- Afterwards: a lot of work on dynamics, abduction, counterfactuals, revision, strong equivalence, etc etc
  - But no similar overview of properties for multiple extensions?
  - NB: If you know some other overview, let me know!

Introduction

- The main question we address is:

  When is a point of view on argument acceptance robust w.r.t. addition/removal of attacks?

- We break this down by considering the following three properties:
  1. XY **addition persistence**: A σ labelling of an AF F in which x is labelled X and y is labelled Y is still a σ labelling of F after adding an attack from x to y.
  2. XY **removal persistence**: A σ labelling of an AF F in which x is labelled X and y is labelled Y is still a σ labelling of F if removing the attack from x to y.
  3. XY **skeptical monotony**: If in all σ labellings of an AF F, x is labelled X and y is labelled Y, then adding an attack from x to y does not lead to new σ labellings.

(Alternative: only consider arguments labeled in.)

Argumentation Frameworks

Definition

Let $\mathcal{U}$ be a set whose elements are called *arguments*. An *argumentation framework* is a pair $F = (A, \rightsquigarrow)$ where A is a finite subset of $\mathcal{U}$ and $\rightsquigarrow \subseteq A \times A$ is a relation called the *attack relation*. We denote by $\mathcal{F}$ the set of all argumentation frameworks.

Labellings

Definition

A *labelling* of an argumentation framework $(A, \rightsquigarrow)$ is a function $L : A \to \{\mathbf{I}, \mathbf{O}, \mathbf{U}\}$. Given a label $l \in \{\mathbf{I}, \mathbf{O}, \mathbf{U}\}$ we define $L^{-1}(l)$ as $\{x \in A \mid L(x) = l\}$. Given an argumentation framework F, we let $\mathcal{L}(F)$ denote the set of all labellings of F.

Definition

Let $F = (A, \rightsquigarrow)$ be an argumentation framework. A labelling $L \in \mathcal{L}(F)$ is *complete* if and only if:

1. for all $x \in A$, $L(x) = \mathbf{I}$ iff for all y s.t. $y \rightsquigarrow x$, $L(y) = \mathbf{O}$.
2. for all $x \in A$, $L(x) = \mathbf{O}$ iff for some y s.t. $y \rightsquigarrow x$, $L(y) = \mathbf{I}$.

Semantics

The complete, grounded, preferred, semi-stable and stable semantics are defined as follows.

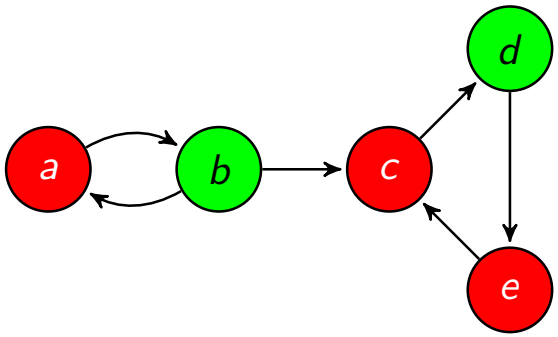
Definition

Let $F = (A, \rightsquigarrow)$ be an argumentation framework.

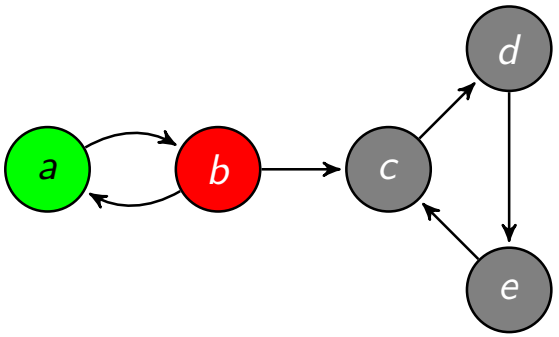
- $\mathcal{L}_{co}(F) = \{L \in \mathcal{L}(F) \mid L \text{ is a complete labelling of } F\}$
- $\mathcal{L}_{gr}(F) = \{L \in \mathcal{L}_{co}(F) \mid \nexists L' \in \mathcal{L}_{co}(F) \text{ s.t.} L'^{-1}(\mathbf{I}) \subset L^{-1}(\mathbf{I})\}$
- $\mathcal{L}_{pr}(F) = \{L \in \mathcal{L}_{co}(F) \mid \nexists L' \in \mathcal{L}_{co}(F) \text{ s.t.} L^{-1}(\mathbf{I}) \subset L'^{-1}(\mathbf{I})\}$
- $\mathcal{L}_{ss}(F) = \{L \in \mathcal{L}_{co}(F) \mid \nexists L' \in \mathcal{L}_{co}(F) \text{ s.t.} L'^{-1}(\mathbf{U}) \subset L^{-1}(\mathbf{U})\}$
- $\mathcal{L}_{st}(F) = \{L \in \mathcal{L}_{co}(F) \mid L^{-1}(\mathbf{U}) = \emptyset\}$

An example

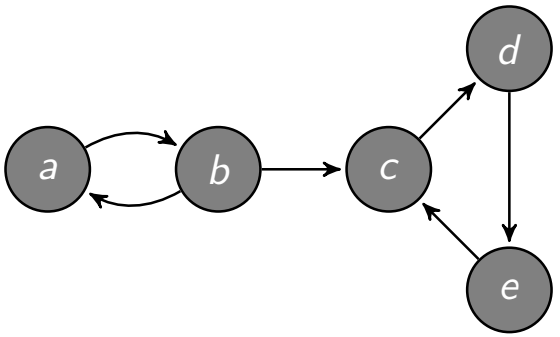
We use green, red and grey to depict in, out and undecided arguments.



: L_1



: L_2



: L_3

	Complete	Grounded	Preferred	Semi-Stable	Stable
L_1	✓		✓	✓	✓
L_2	✓		✓		
L_3	✓	✓			

Addition Persistence Properties

Definition (XY Addition Persistence)

A semantics σ satisfies XY addition persistence iff every σ labelling of an AF F in which x is labelled X and y is labelled Y is still a σ labelling of F after adding an attack from x to y.

Note: only grounded, complete, preferred and semi-stable semantics.

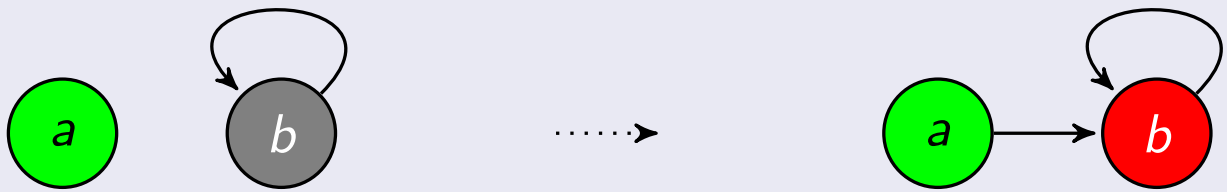
We discuss stable semantics later (roughly: trivial or same as preferred).

Addition Persistence Properties

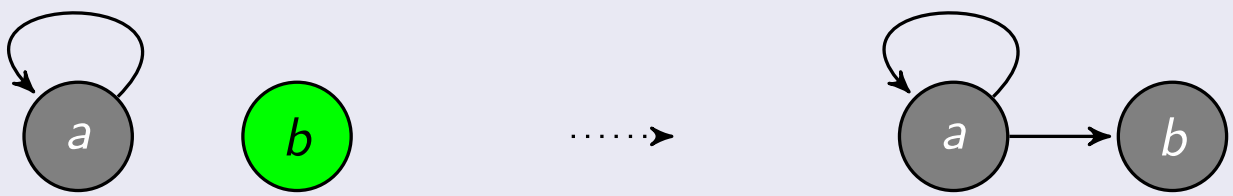
Some addition persistence properties fail simply because they introduce a conflict. The following examples apply to all semantics we consider.

Failure of **II**-addition persistence

Failure of **IU**-addition persistence



Failure of **UI**-addition persistence



Addition Persistence Properties

As we just saw, some properties fail:

- **II**-addition persistence
- **IU**-addition persistence
- **UI**-addition persistence

} These cases fail because they introduce a conflict.

Other properties reflect reasonable principles:

- **OO**-addition persistence
- **OU**-addition persistence
- **OI**-addition persistence
- **IO**-addition persistence
- **UO**-addition persistence
- **UU**-addition persistence

} In these cases the added attack doesn't introduce a conflict, and doesn't invalidate the justification of the attacked argument's label.

Are these properties satisfied by the semantics we consider?

Addition Persistence Properties

Grounded:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	✓	-
	I	-	-	-

Complete:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	✓	-
	I	✓	-	-

Preferred:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	-	-
	I	✓	-	-

Semi-Stable:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	-
	I	-	-	-

Addition Persistence Properties

Grounded:

		X		
		O	U	I
Y	O	✓	✓	✓
	U	✓	✓	-
	I	-	-	-

Complete:

		X		
		O	U	I
Y	O	✓	✓	✓
	U	✓	✓	-
	I	✓	-	-

Preferred:

		X		
		O	U	I
Y	O	✓	✓	✓
	U	✓	-	-
	I	✓	-	-

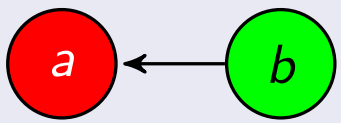
Semi-Stable:

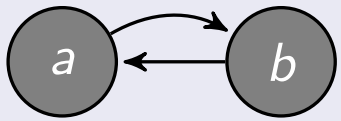
		X		
		O	U	I
Y	O	-	-	-
	U	-	-	-
	I	-	-	-

Addition Persistence Properties

Failure of **OI** addition persistence under grounded semantics.

After adding an attack from a to b there is a new grounded labelling:





Addition Persistence Properties

Grounded:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	✓	-
	I	-	-	-

Complete:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	✓	-
	I	✓	-	-

Preferred:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	-	-
	I	✓	-	-

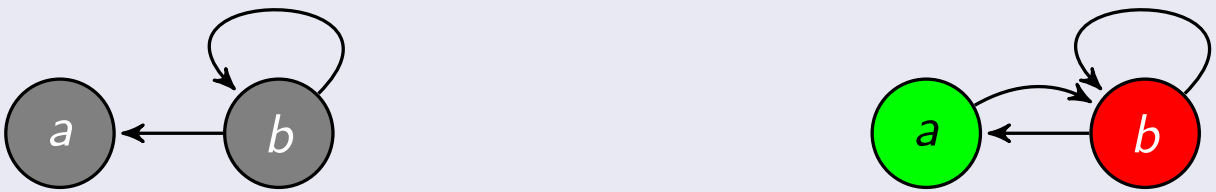
Semi-Stable:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	-
	I	-	-	-

Addition Persistence Properties

Failure of UU addition persistence under the preferred semantics.

After adding an attack from a to b there is a new preferred labelling:



Addition Persistence Properties

Grounded:

		X		
		O	U	I
Y	O	✓	✓	✓
	U	✓	✓	-
	I	-	-	-

Complete:

		X		
		O	U	I
Y	O	✓	✓	✓
	U	✓	✓	-
	I	✓	-	-

Preferred:

		X		
		O	U	I
Y	O	✓	✓	✓
	U	✓	-	-
	I	✓	-	-

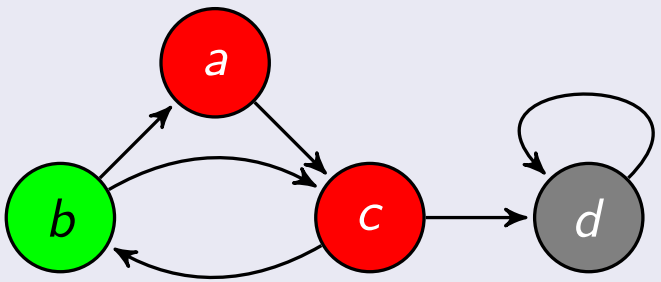
Semi-Stable:

		X		
		O	U	I
Y	O	-	-	-
	U	-	-	-
	I	-	-	-

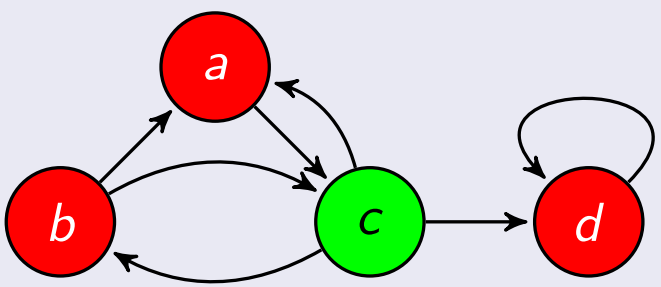
Addition Persistence Properties

Failure of **OO** addition persistence under the semi-stable semantics

The following AF has one semi-stable labelling:



The arguments a and c are both **O**. But if we add an attack $c \rightsquigarrow a$ we get a different semi-stable labelling:



Removal Persistence Properties

Definition (XY Removal Persistence)

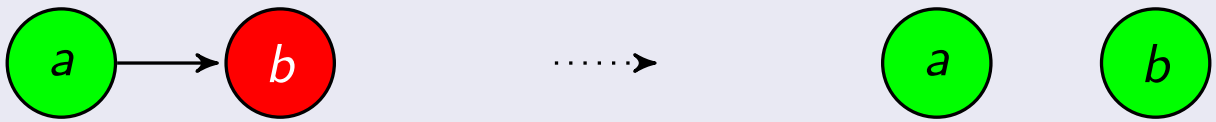
A semantics σ satisfies XY removal persistence iff: every σ labelling of an AF F in which x is labelled X and y is labelled Y is still a σ labelling of F after removing the attack from x to y.

Note: again, we look only at the grounded, complete, preferred and semi-stable semantics. We discuss the stable semantics later.

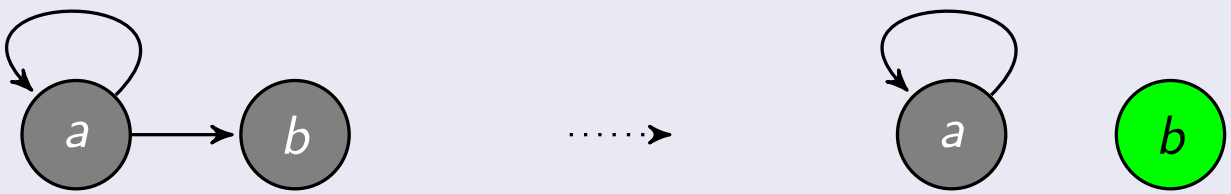
Removal Persistence Properties

Some removal persistence properties fail because the removal of an attack from x to y may make the label of y unjustified. The following examples apply for all semantics we consider.

Failure of **IO**-removal persistence:



Failure of **UU**-removal persistence:



Removal Persistence Properties

As we just saw, some removal persistence properties fail:

- **IO**-removal persistence
- **UU**-removal persistence

} These cases fail because they invalidate the justification of the attacked argument's label.

Others are trivially satisfied:

- **II**-removal persistence
- **UI**-removal persistence
- **IU**-removal persistence

} Combinations of labels that never occur.

The remaining properties reflect reasonable principles:

- **OO**-removal persistence
- **OU**-removal persistence
- **OI**-removal persistence
- **UO**-removal persistence

} In these cases the removal does not invalidate the attacked argument's label.

Are these properties satisfied by the semantics we consider?

Removal Persistence Properties

Grounded:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	-
	U	✓	-	✓
	I	✓	✓	✓

Complete:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	-
	U	✓	-	✓
	I	✓	✓	✓

Preferred:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	-
	U	✓	-	✓
	I	✓	✓	✓

Semi-Stable:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	✓
	I	-	✓	✓

Removal Persistence Properties

Grounded:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	-
	U	✓	-	✓
	I	✓	✓	✓

Complete:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	-
	U	✓	-	✓
	I	✓	✓	✓

Preferred:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	-
	U	✓	-	✓
	I	✓	✓	✓

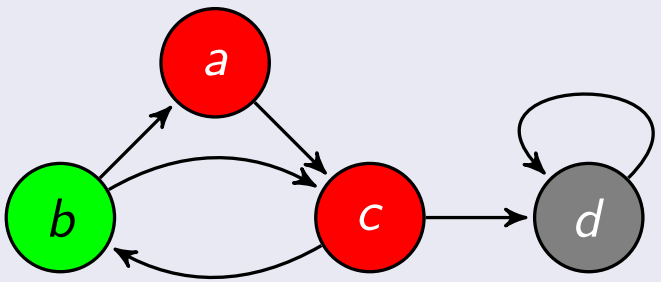
Semi-Stable:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	✓
	I	-	✓	✓

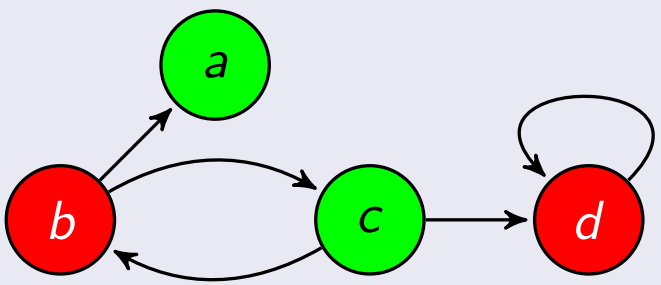
Removal Persistence Properties

Failure of **OO** removal persistence under the semi-stable semantics

The following AF has one semi-stable labelling:



The arguments a and c are both **O**. But if we remove the attack $a \rightsquigarrow c$ we get a different semi-stable labelling:



Removal Persistence Properties

Grounded:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	-
	U	✓	-	✓
	I	✓	✓	✓

Complete:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	-
	U	✓	-	✓
	I	✓	✓	✓

Preferred:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	-
	U	✓	-	✓
	I	✓	✓	✓

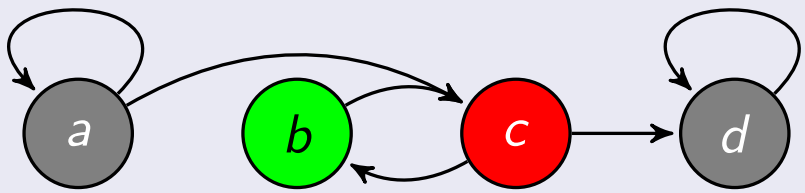
Semi-Stable:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	✓
	I	-	✓	✓

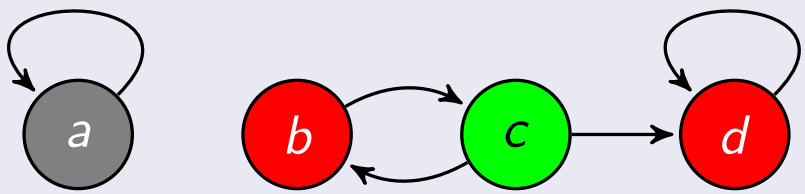
Removal Persistence Properties

Failure of **UO** removal persistence under the semi-stable semantics

The following AF has one semi-stable labelling:



The arguments a and c are labelled **U** and **O**. If we remove the attack $a \rightsquigarrow c$ we get a different semi-stable labelling:



Removal Persistence Properties

Grounded:

		X		
		O	U	I
Y	O	✓	✓	-
	U	✓	-	✓
	I	✓	✓	✓

Complete:

		X		
		O	U	I
Y	O	✓	✓	-
	U	✓	-	✓
	I	✓	✓	✓

Preferred:

		X		
		O	U	I
Y	O	✓	✓	-
	U	✓	-	✓
	I	✓	✓	✓

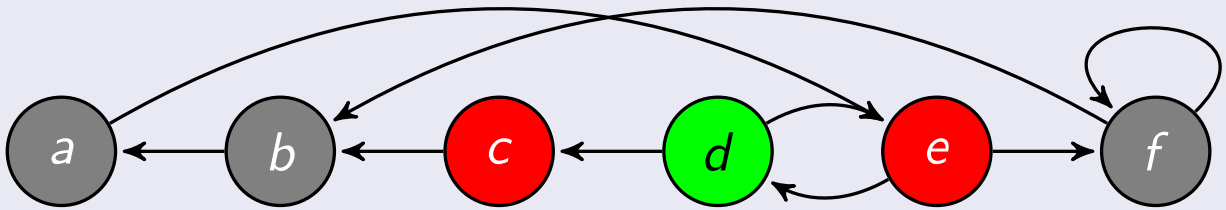
Semi-Stable:

		X		
		O	U	I
Y	O		-	-
	U	-	-	✓
	I	-	✓	✓

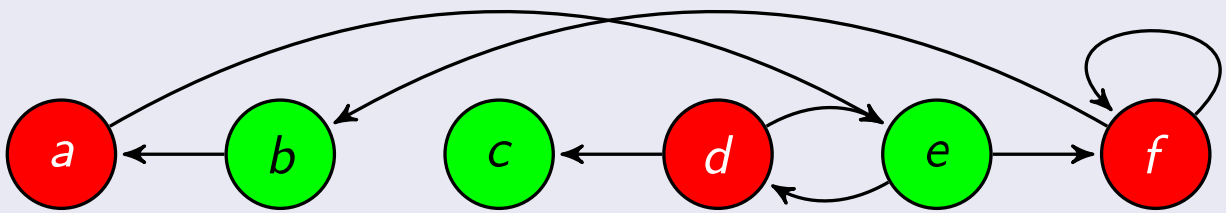
Removal Persistence Properties

Failure of **OU** removal persistence under the semi-stable semantics

The following AF has one semi-stable labelling:



The arguments c and b are labelled **O** and **U**. If we remove the attack $c \rightsquigarrow b$ we get a different semi-stable labelling:



Skeptical Monotony Properties

Intuition

Suppose that x and y are labelled X and Y in *all* σ labellings of F. The XY addition persistence property then implies that all σ labellings of F are still σ labellings after adding the attack $x \rightsquigarrow y$. In other words, *no* σ *labelling gets destroyed*. But is it also the case that *no new labellings are created?* This is the property that we consider here.

Definition (Skeptical XY monotony)

A semantics σ satisfies skeptical XY monotony iff: If in all σ labellings of an AF F, x is labelled X and y is labelled Y, then adding an attack from x to y does not lead to new σ labellings.

Note: again, we look only at the grounded, complete, preferred and semi-stable semantics. We discuss the stable semantics later.

Skeptical Monotony Properties

Some skeptical monotony properties fail simply because they introduce a conflict and thus lead to new labellings. The following examples apply to all semantics we consider.

Failure of **II** skeptical monotony

Failure of **IU** skeptical monotony

Failure of **UI** skeptical monotony

Skeptical Monotony Properties

The previous examples show that the following properties fail:

- **II** skeptical monotony
- **IU** skeptical monotony
- **UI** skeptical monotony

} These cases fail because they introduce a conflict and thus produce new labellings.

What about the other properties?

- **OO** skeptical monotony
- **OU** skeptical monotony
- **OI** skeptical monotony
- **IO** skeptical monotony
- **UO** skeptical monotony
- **UU** skeptical monotony

Addition Persistence Properties (reminder)

Grounded:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	✓	-
	I	-	-	-

Complete:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	✓	-
	I	✓	-	-

Preferred:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	-	-
	I	✓	-	-

Semi-Stable:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	-
	I	-	-	-

Skeptical Monotony Properties

Grounded:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	✓	-
	I	-	-	-

Complete:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	-	-
	I	-	-	-

Preferred:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	-
	I	-	-	-

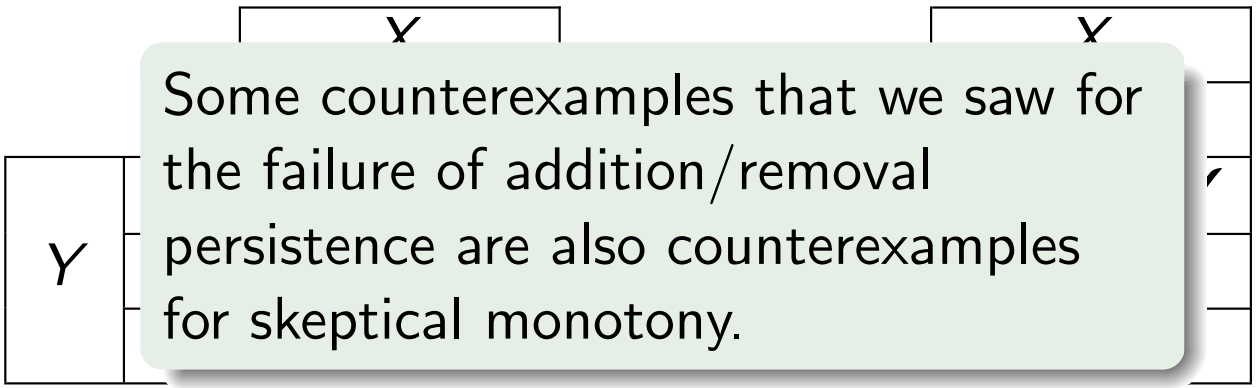
Semi-Stable:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	-
	I	-	-	-

Skeptical Monotony Properties

Grounded:

Complete:



Some counterexamples that we saw for the failure of addition/removal persistence are also counterexamples for skeptical monotony.

Preferred:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	-
	I	-	-	-

Semi-Stable:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	-
	I	-	-	-

Skeptical Monotony Properties

Grounded:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	✓	-
	I	-	-	-

Complete:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	-	-
	I	-	-	-

Preferred:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	-
	I	-	-	-

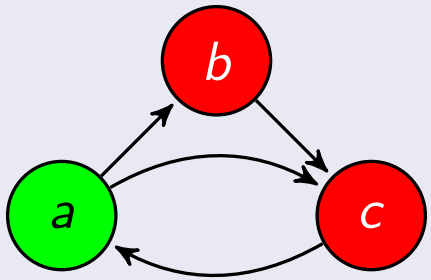
Semi-Stable:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	-
	I	-	-	-

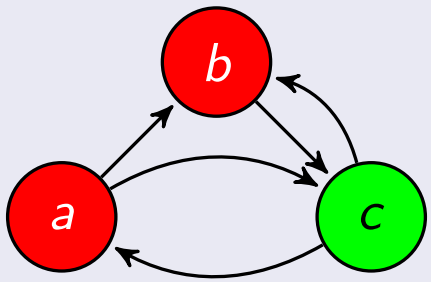
Skeptical Monotony Properties

Failure of **OO** skeptical monotony under the preferred semantics

The following AF has one preferred labelling:



The arguments b and c are both labelled **O**. If we add an attack $c \rightsquigarrow b$ then we obtain a new preferred labelling:



Skeptical Monotony Properties

Grounded:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	✓	-
	I	-	-	-

Complete:

		X		
		O	**U**	**I**
Y	**O**	✓	✓	✓
	U	✓	-	-
	I	-	-	-

Preferred:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	-
	I	-	-	-

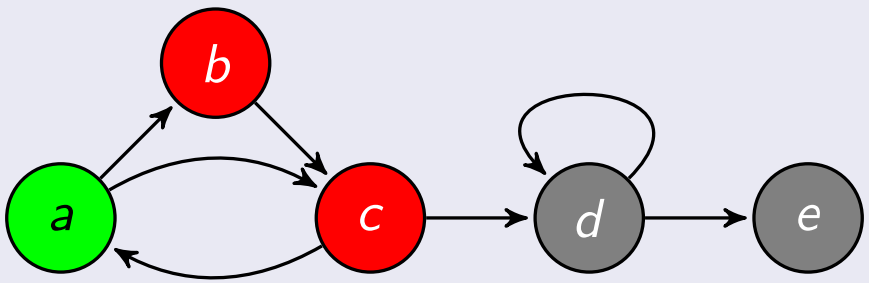
Semi-Stable:

		X		
		O	**U**	**I**
Y	**O**	-	-	-
	U	-	-	-
	I	-	-	-

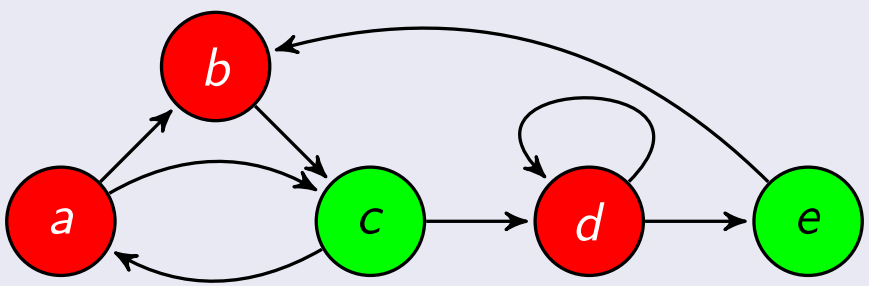
Skeptical Monotony Properties

Failure of **UO** skeptical monotony under the preferred semantics

The following AF has one preferred labelling:



The arguments e and b are labelled **U** and **O**. If we add an attack $e \rightsquigarrow b$ then we obtain a new preferred labelling:



Some Remarks About the Stable Semantics

- All properties that we considered that involve **U** labelled arguments are trivially satisfied under the stable semantics, because no argument is ever labelled **U** under the stable semantics.
- For all properties that do not involve **U** labelled arguments, the results for the stable semantics are *exactly the same as for the preferred semantics*.

Conclusions

- We studied the behaviour of semantics for argumentation when the AF changes.
- Some insights we gained are that some properties that appear intuitive are not satisfied under all semantics
  - Failure of e.g. **OO** skeptical monotony under the preferred semantics.
  - Failure of *many* properties under the semi-stable semantics.
- The complete semantics is the best behaved semantics w.r.t. the properties we discussed.

Future Research

- Study skeptical removal monotony.
- Study weaker versions of the properties considered here.
- Study semantics not considered here.
- Study relationships with other work on the behaviour of semantics:
  - Strong equivalence.
  - Input/output behaviour.
  - The directionality principle.
- Study the consequences of our results in the setting of strategic argumentation and revision, counterfactuals and abduction in argumentation.

Dagstuhl workshop